Artificial Intelligence (521495A), Spring 2022
Exercise 4 : Machine learning
Deadline for reports: Wed 16.2.2022 23:59 (+1h)

This handout contains five problems related to machine learning part of the course (in lecture track 2). Problems 1-4 are pre-exercises to support learning and solutions to them are provided in **solutions4.pdf**. For Problem 5, the solution is not given and you should return your solution as a report. **Include in the report only the solution to Problem 5**. The solution gives max. 1 point, which is taken into account in course grading.

Problems 1-5: Background material is provided in Lectures 7 and 9 (Track 2). See also the course book Chapters 18 and 20.

Problem 1. Consider a dataset below where four binary features (i.e., input variables) are presented as f_1, f_2, f_3, and f_4, as well as corresponding binary class label y.

f_1	f_2	f_3	f_4	y
0	1	1	1	0
1	1	0	1	0
1	0	0	0	1
0	0	1	1	1
0	1	0	0	0
0	1	1	0	0

(a) Draw a Bayesian network for this problem (i.e., naive Bayes model).

(b) Calculate maximum likelihood estimate of $P(f_1 \mid y = 0)$.

(c) Calculate Laplace smoothed estimate of $P(f_3 \mid y = 1)$, with $k = 1$.

(d) Calculate Laplace smoothed estimate of $P(y = 1)$, with $k = 3$.

Problem 2. Football team manager has collected a labeled data set to be able to predict future game results. Target label of game result is y (+1: win, 0: tie, -1: lose) for the given input feature vector $\mathbf{x} = (h, w, t)$. There are three binary input features (1 or 0): h (home game or not), w (won the previous game or not), and t (team is in good shape or not).

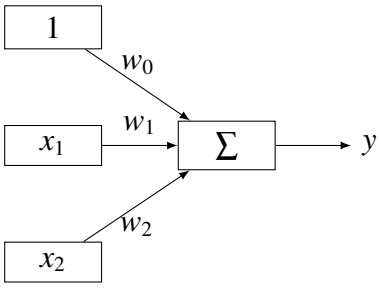
h	w	t	y
1	1	1	+1
0	1	1	+1
1	0	1	0
0	1	0	0
0	0	1	-1
0	0	0	-1

(a) To predict future game results, manager is building a Naive Bayes (NB) classifier. Form and fill the model parameter tables for prior and conditional probabilities based on training data set above.

(b) Form a prediction equation for this NB classifier and calculate the prediction of the result of upcoming home game when team has win the previous game but has injured players (i.e., feature vector is $\mathbf{x} = (1, 1, 0)$). What will be the predicted result?

(c) Use Laplace smoothing with $k = 1$. What will be the predicted game result now?

Problem 3. Consider the following perceptron model, for which the inputs are the constant 1 feature (bias) and two binary features $x_1 \in \{0, 1\}$ and $x_2 \in \{0, 1\}$. The output is $y \in \{0, 1\}$.



(a) Show visually if there is a proper solution (i.e., weight vector) for the logical XOR and AND operations using perceptron model. (XOR is the logical exclusive OR operation, which equals to zero when x_1 equals to x_2 and equals to one when x_1 is different from x_2).

(b) Show what conditions of weight vector satisfy the logical AND operation.

Problem 4. Consider a dataset below with two continuous features f_1 and f_2, taking values between 0-2 (i.e., $f_1 \in \mathbb{R}|0 \leq f_1 \leq 2$ and $f_2 \in \mathbb{R}|0 \leq f_2 \leq 2$), as well as corresponding discrete class label $y \in \{-1, 0, +1\}$.

f_1	f_2	y
0	0	-1
0	1	-1
1	1	0
2	1	0
1	0	+1
2	2	+1

(a) Show visually if classes are linearly separable.

(b) An unknown example $\mathbf{x} = (f_1, f_2) = (1.5, 2)$ is given. What is the predicted class using k-nearest neighbor classifier with $k = 3$? (show visually)

Problem 5* [1p]. Consider the task of building a model for water quality estimation using sensor that is able to measure turbidity t (i.e., how clear the water is) and microbe content m of the water. The sensor unit gives continuous values from 0 to 1 for both variable so that 0 is very low and 1 is very high turbidity level and microbe content, respectively. Field expert has collected a data set and water examples are labeled in laboratory with labels y (+1: good quality, 0: normal quality, -1: bad quality) for each input feature vector $\mathbf{x} = (t, m)$. Dataset is presented below.

bias	t	m	y
1.0	1.0	1.0	-1
1.0	0.25	0.5	+1
1.0	0.75	1.0	-1
1.0	1.0	0.5	0
1.0	0.75	0.0	0
1.0	0.0	0.0	+1

(a) To predict the quality of unknown water samples, you task is to train a perceptron classifier. Run perceptron training algorithm and form a table of the updated weights after first and second data example from the dataset above. Weight vectors are initialised as $\mathbf{w}_{y=-1} = (0,0,0)$, $\mathbf{w}_{y=0} = (1,0,0)$, and $\mathbf{w}_{y=+1} = (0,0,0)$. Calculate the weights and fill the table below.

	$\mathbf{w}_{y=-1}$			$\mathbf{w}_{y=0}$			$\mathbf{w}_{y=+1}$		
Initial weights	0	0	0	1	0	0	0	0	0
Training: $(1.0, 1.0, 1.0) \rightarrow -1$									
Training: $(1.0, 0.25, 0.5) \rightarrow +1$									

(b) Show visually if the given dataset above is linearly separable or not (with this multiclass setting).

(c) Write down the perceptron prediction equation and predict the water quality for $\mathbf{x} = (1.0, 0.75, 0.25)$, when the learned weight vectors are $\mathbf{w}_{y=-1} = (1.0, 1.0, -1.0)$, $\mathbf{w}_{y=0} = (1.0, 2.0, -2.0)$, and $\mathbf{w}_{y=+1} = (-1.0, -1.0, 1.0)$. What will be the predicted water quality class?